REMARKS

Applicants have studied the final Office Action dated December 5, 2000, and respectfully request consideration of this response under the provisions of 37 C.F.R. § 1.116 in that remarks below establish that the claims are in condition for allowance. Claims 1-27 are pending, and no claims are requested to be amended by this amendment.

Claims 1, 3, 5, 8, 9, 11, 13, 16-20 and 24-27 were rejected under 35 U.S.C. §102(e) as being anticipated by Kageyama (U.S. Patent No. 5,955,693). Claims 4, 14 and 22 were rejected under 35 U.S.C. §103(a) as being unpatentable over Kageyama. These rejections are respectfully traversed.

The present invention is directed to a voice converter that allows imitations of professional singers to be performed. In one embodiment, the voice converter includes an extracting means for extracting a plurality of sinusoidal wave components from an input voice signal, including frequencies and/or amplitudes of the sinusoidal wave components. Here, the plurality of sinusoidal wave components, including frequencies and/or amplitudes of the sinusoidal wave components, are spectral wave components of the input voice. For example, the set of sinusoidal wave components, representing the spectral wave components of the input voice, may be in the form of frequency value F and amplitude value A coordinates, such as (F0, A0), (F1, A1), (F2, A2), ...(Fn, An) (see Figs. 2 and 3 of the current application). In one exemplary implementation, as shown in Fig. 1 of the current application, a microphone 1 gathers a karaoke singer's voice and provides an input voice signal Sv. The input voice signal Sv is then analyzed by a Fast Fourier Transform (FFT) section 2, and the frequency spectrum thereof is detected (see page 8, lines 12-16 of the current application). The processing implemented by the FFT section 2 is carried out in prescribed frame units, so that a frequency spectrum is created successively for each frame (see Fig. 2 of the current application). A peak detecting section 3 detects peaks in the frequency spectrum of the input voice signal Sv. For example, sampling results of the plurality of sinusoidal wave components Fn, An, as illustrated in part (1) of Fig. 6

of the current application, are obtained. The frequency Fn and the amplitude An of each sinusoidal component are modified according to a pitch and volume of a model voice signal, as depicted in Fig. 6 of the current application. Accordingly, when the voice of the karaoke singer is outputted after the modification, the characteristics of the voice, the manner of singing, and the like are significantly influenced by the model voice signal.

As grounds for the rejection of claims 1, 3, 5, 8, 9, 11, 13, 16-20 and 24-27 in the final Office Action dated December 5, 2000, the Examiner contended that the aforementioned "pair of the original frequency (Fn) and the original amplitude (An) representing a corresponding sinusoidal wave component [contained in the input voice signal]" is disclosed in Kageyama. In support of this contention, the Examiner cited column 6, lines 7-14, column 6, lines 61-63 and column 7, lines 1-5 (see Page 3 of Examiner's Final Rejection). Applicants respectfully disagree with the Examiner's contention. It is respectfully submitted that Kageyama discloses a karaoke apparatus which differs from that which is presently claimed.

Kageyama discloses a karaoke apparatus and method for modifying a live singing voice (*i.e.,* input voice) to a voice similar to the original/model singer (*i.e.*, model voice) of the karaoke song. However, the voice modifying method employed in Kageyama's karaoke apparatus is different from the present voice conversion technique which uses the set of sinusoidal components Fn, An extracted from an input voice signal, <u>the set of sinusoidal components representing spectral wave components of the input voice signal</u>. Kageyama's karaoke apparatus takes an input voice signal (or live singing voice) and separates it into a lead consonant component and a subsequent vowel component. From the subsequent vowel component, a secondary characteristic, which may for example be the pitch of the component, is extracted. A primary characteristic of a corresponding model vowel contained in the model voice is also supplied. The primary characteristic of the model vowel is represented by phoneme data, in terms of the waveform, envelope thereof, vibrato frequency, vibrato depth and supplemental noise (see column 4, line 66 to column 5, line 2 of the Kageyama reference).[1] A substitutive

---

[1] Referring to Fig. 6A of the Kageyama reference, a phrase of lyric "A KA SHI YA NO" comprises five syllables "A", "KA", "SHI", "YA" and "NO", and the phoneme data are composed of extracted vowels "a", "a", "I", "a" and "o" from the five syllables.

vowel component is then created according to the primary characteristics of the model vowel and the secondary characteristics (*e.g.*, pitch) of the subsequent vowel component of the input voice. Finally, the substitutive vowel component is combined with the lead consonant component to synthesize an output singing voice.

With respect to column 6, lines 7-14 cited by the Examiner, the vowel component of the input voice signal (or "singing voice signal") referred to in the Kageyama reference is different from the set of sinusoidal components Fn, An of the present invention. The paragraph cited by the Examiner reads:

> A microphone 27 constitutes an input device and collects or picks up a singing voice signal, which is fed to the voice converter DSP 30 through a pre-amplifier 28 and an A/D converter 29. The DSP 30 converts each vowel component of the singing voice signal into a substitutive vowel component which is created according to a vowel waveform of a model person such as an original singer. The converted signal is put into the sound effect DSP 20.

The "vowel component" referred in Kageyama refers to a vowel portion of a syllable in a phrase of lyric (see column 4, lines 57-66 of the Kageyama reference). In Fig. 6A of the Kageyama reference, a phrase of lyric 'A KA SHI YA NO' comprises five syllables 'A', 'KA', 'SHI', 'YA', 'NO.' Every syllable has a vowel component and/or a consonant component. For example, in the third syllable 'SHI', the vowel component is 'I' and the consonant component is 'SH'. This vowel component is replaced with the substitute vowel component, which is created according to primary characteristics of a corresponding model vowel and secondary characteristics of the vowel component. The vowel portion of a syllable in an input voice signal, as described by Kageyama, is different from the set of sinusoidal wave components Fn, An from the input voice signal as in the present invention. In the present invention, the set of sinusoidal wave components extracted from the input voice signal are <u>spectral wave components of the input voice signal</u>, and not a vowel portion of a syllable in the input voice signal.

With respect to column 6, lines 61-63 cited by the Examiner, the leading consonant component and the subsequent vowel component of a syllable contained in the input voice signal referred to in the Kageyama reference are different from the set of sinusoidal components Fn, An

disclosed in the present invention. The portion cited by the Examiner is in a paragraph that reads:

> A consonant separator 40 accepts a digitized input singing voice signal collected through the microphone 27, the pre-amplifier 28, and the A/D converter 29. The consonant separator 40 separates a leading consonant component and a subsequent vowel component of each syllable contained in the digitized input singing voice signal.....

The "leading consonant component" and the "subsequent vowel component" in Kageyama refer to a consonant portion and vowel portion of a syllable in a phrase of lyric, respectively (see column 4, lines 57-66 of the Kageyama reference). When a phrase of lyric is inputted through a microphone 27 in Kageyama's karaoke apparatus, it is first digitized, creating a digitized input singing voice. Digitized consonant component and digitized vowel component of a particular syllable are then separated. A substitute vowel component, created according to primary characteristics of a model vowel and the secondary characteristics of the subsequent vowel component, replaces the subsequent vowel component. This modifies the input singing voice signal to a voice signal that is similar to the model singer. However, the digitized consonant portion and the digitized vowel portion of a syllable in an input singing voice signal, as described by Kageyama, are different from the set of sinusoidal wave components Fn, An extracted from an input voice signal in the present invention. As described above, the set of sinusoidal wave components extracted from the input voice signal are <u>spectral wave components of the input voice signal</u>, and not a digitized consonant portion and a digitized vowel portion of a syllable in an input voice signal.

With respect to column 7, lines 1-5 cited by the Examiner, the pitch (frequency) and the level of the digitized vowel portion of a syllable contained in the input voice signal referred to in the Kageyama reference are different from the set of sinusoidal components Fn, An of the present invention. The portion cited by the Examiner reads:

> The pitch/level detector 41 constitutes an analyzing device to analyze the input singing voice signal to extract therefrom secondary characteristics. Namely, the detector 41 detects the pitch (frequency) and the level of the input vowel component.

The "pitch" mentioned in the above portion of the Kageyama reference refers to the fundamental frequency of the digitized vowel portion of a syllable contained in the input singing voice signal. This is not the same as frequencies Fn of the present invention, which is a frequency of each sinusoidal wave component in the frequency spectral wave components F1-Fn of an input voice signal. The "level" mentioned in the above portion of the Kageyama reference refers to the volume or envelop of the digitized vowel portion of a syllable contained in the input singing voice signal. This is not the same as amplitude An disclosed in the present invention, which is an amplitude of each sinusoidal wave component in the amplitude spectral wave components A1-An of an input voice signal.

The karaoke apparatus of Kageyama does not disclose, teach or suggest extracting a set of sinusoidal wave components representing spectral wave components of an input voice signal, the sinusoidal wave components including frequencies Fn of the sinusoidal wave components and/or amplitudes An of the sinusoidal wave components, and modulating these component, as recited in claims 1, 9, 25 and 26. Likewise, the karaoke apparatus of Kageyama does not disclose, teach or suggest an analyzer device that analyzes a plurality of sinusoidal wave components contained in the input voice signal to derive a parameter set of an original frequency and an original amplitude, each pair of the original frequency and the original amplitude representing a corresponding sinusoidal wave component, and a modulator device that modulates the parameter set of the sinusoidal wave components according to reference information, as recited in claims 17 and 27. Therefore, it is respectfully submitted that claims 1, 9, 17 and 25-27 distinguish over the Kageyama reference. Because each dependent claim incorporates all the limitations of its base claim(s), claims depending from 1, 9 and 17 also distinguish over the Kageyama reference. Claims 3, 5 and 8 depend directly or indirectly from claim 1. Claims 11, 13 and 16 depend directly or indirectly from claim 9. Claims 18-20 and 24 depend directly or indirectly from claim 17. Accordingly, it is respectfully submitted that the rejections of claims 1, 3, 5, 8, 9, 11, 13, 16-20 and 24-27 under 35 U.S.C. §102(e) and claims 4, 14 and 22 under 35 U.S.C. §103(a) should be withdrawn.

Claims 2, 6, 10, 12 and 21 were rejected under 35 U.S.C. §103(a) as being unpatentable over Kageyama in view of Matsumoto '303 (U.S. Patent No. 5,847,303). Claims 7, 15 and 23 were rejected under 35 U.S.C. §103(a) as being unpatentable over Kageyama in view of Matsumoto '907 (U.S. Patent No. 5,963,907). These rejections are respectfully traversed.

The claimed features of the present invention are not realized even if the teachings of the Matsumoto '303 reference or Matsumoto '907 reference are incorporated into Kageyama. Matsumoto '303 is directed to a voice processing apparatus that modulates an input voice signal into an output voice signal according to a set of parameters. Matsumoto '303 discloses a voice change parameter table of filter coefficients to control spectrum shape of varying pitch ranges for the purpose of providing more realistic sounding conversion between male and female voices (see Figs. 9 and 10; column 11, lines 3-26 of the Matsumoto '303 reference). An audio signal processor within the voice processing apparatus is configured by a parameter set to process the audio signal by modifying the frequency spectrum of the input voice. However, Matsumoto '303 does not disclose the inventive features of the present invention in extracting a plurality of sinusoidal wave components from an input voice signal <u>representing frequency spectral wave components of the input voice signal</u>, the sinusoidal wave components including frequencies of the sinusoidal wave components and modulating frequencies of the sinusoidal wave components according to pitch information representative of a pitch of a reference voice signal, as is recited in claim 1. Likewise, Matsumoto '303 does not disclose extracting a plurality of sinusoidal wave components from the input voice signal <u>representing amplitude spectral wave components of the input voice signal</u>, the sinusoidal wave components including amplitudes of the sinusoidal wave components and modulating amplitudes of the sinusoidal wave component extracted from the input voice signal according to the amplitude information representative of amplitudes of sinusoidal wave components contained in a reference voice signal, as is recited in claim 9. Claim 17 incorporates the above limitations of claims 1 and 9; therefore, it also distinguishes over the Matsumoto '303 reference.

Matsumoto '907 is directed to a voice converter that provides pitch and formant shifting of an input voice signal. Referring to Fig. 2 of the Matsumoto '907 reference, an audio filter 325

extracts the volume level of the input voice signal, and outputs the extracted volume level as first volume data V1. A second audio filter 326 extracts the volume level of an output voice signal, and outputs the extracted volume level as second volume data V2. A difference judging circuit 322 compares the first and second volume data V1 and V2 with each other, and determines a volume gain G and a distorting factor D which is supplied to a distortion circuit 321. When the volume of the output voice after conversion is smaller than that of the input voice, the volume gain G is increased. In contrast, the subject matter of claims 7, 15 and 23 in the present invention is to change the volume of an input singing voice in matching with the variation of the volume of the voice of a model singer. This allows the volume of an output voice signal to emulate the volume variation of the reference voice signal of the model singer. Such feature is not disclosed, taught or suggested by Matsumoto '907. Additionally, Matsumoto '907 does not disclose the inventive features of the present invention in extracting a plurality of sinusoidal wave components from an input voice signal representing frequency spectral wave components of the input voice signal, the sinusoidal wave components including frequencies of the sinusoidal wave components and modulating frequencies of the sinusoidal wave components according to pitch information representative of a pitch of a reference voice signal, as is recited in claim 1. Likewise, Matsumoto '907 does not disclose extracting a plurality of sinusoidal wave components from the input voice signal representing amplitude spectral wave components of the input voice signal, the sinusoidal wave components including amplitudes of the sinusoidal wave components and modulating amplitudes of the sinusoidal wave component extracted from the input voice signal according to the amplitude information representative of amplitudes of sinusoidal wave components contained in a reference voice signal, as is recited in claim 9. Claim 17 incorporates the above limitations of claims 1 and 9; therefore, it distinguishes over the Matsumoto '907 reference.

Applicant believes that the differences between Kageyama, Matsumoto '303, Matsumoto '907 and the present invention are clear in claims 1, 9 and 17, which set forth voice conversion and synthesizing apparatuses that utilize a plurality of sinusoidal wave components according to embodiments of the present invention. Therefore, claims 1, 9 and 17 distinguish over the Kageyama, Matsumoto '303 and Matsumoto '907 references. Claims depending directly or

indirectly from claims 1, 9 and 17, such as claims 2, 6, 7, 10, 12, 15, 21 and 23, also distinguish over the above references. Applicant further believes that the differences between Kageyama, Matsumoto '907 and the present invention are clear in claims 7, 15 and 23, which set forth apparatuses that emulate volume variation of a model singer according to embodiments of the present invention. Therefore, the rejection of claims 2, 6, 7, 10, 12, 15, 21 and 23 under 35 U.S.C. § 103(a) should be withdrawn.

In view of the foregoing, it is respectfully submitted that the application and all of the claims are in condition for allowance. Reexamination and reconsideration of the application are requested.

If for any reason the Examiner finds the application other than in condition for allowance, the Examiner is invited to call the undersigned attorney at (213) 488-7100 should the Examiner believe a telephone interview would advance the prosecution of the application.

Respectfully submitted,

March 1, 2001

By: Roger R. Wise

Roger R. Wise
Reg. No. 31,204

Pillsbury Madison & Sutro LLP
725 South Figueroa Street, Suite 2800
Los Angeles, CA 90017-5406
Telephone (213) 488-7100; Facsimile (213) 629-1033